Interpreting the universal phylogenetic tree

Carl R. Woese*

Department of Microbiology, University of Illinois at Urbana-Champaign, B103 Chemical and Life Sciences Laboratory, MC-110, 601 South Goodwin Avenue, Urbana, IL 61801-3709

Contributed by Carl R. Woese, May 22, 2000

The universal phylogenetic tree not only spans all extant life, but its root and earliest branchings represent stages in the evolutionary process before modern cell types had come into being. The evolution of the cell is an interplay between vertically derived and horizontally acquired variation. Primitive cellular entities were necessarily simpler and more modular in design than are modern cells. Consequently, horizontal gene transfer early on was pervasive, dominating the evolutionary dynamic. The root of the universal phylogenetic tree represents the first stage in cellular evolution when the evolving cell became sufficiently integrated and stable to the erosive effects of horizontal gene transfer that true organismal lineages could exist.

Archaea | Bacteria | Eucarya | universal ancestor | horizontal gene transfer

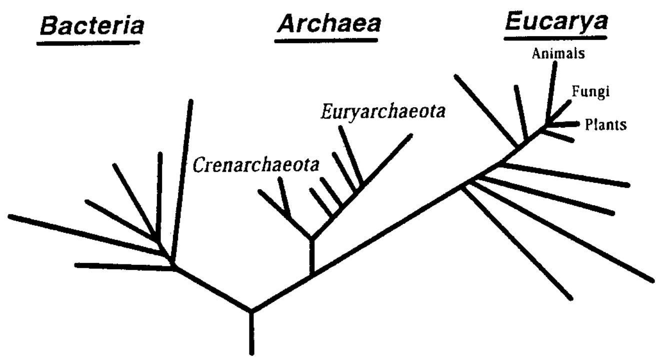

Fig. 1. The basal universal phylogenetic tree inferred from comparative analyses of rRNA sequences (4, 5). The root has been determined by using the paralogous gene couple EF-Tu/EFG (6).

The Grand Challenge

In a letter to T. H. Huxley in 1857, Darwin, with characteristic prescience, foresaw "[t]he time . . . when we shall have very fairly true genealogical trees of each great kingdom of nature" (1), voicing in the terms of his day one of the great, defining challenges of Biology. Another century would pass, however, before Darwin's vision became reality. Darwin obviously knew that the methodologies of the day, paleontology and classical taxonomy, were not up to a task this monumental. What could not be foreseen, however, was that, as Biology moved to a molecular footing in the following century, evolution would cease to be a focus, and what Darwin considered a basic problem would effectively fade from view. Yet a vision this central, this essentially biological, cannot remain forever obscured. In the 1960s, with the advent of molecular sequencing, gene histories and organismal genealogies emerged on the molecular stage (2); and with the recent eruption of genomic sequencing, the full history of cellular life on this planet seems now to be unfolding before our eyes.

What molecular sequences taught us in the 1960s was that the genealogical history of an organism is written to one extent or another into the sequences of each of its genes, an insight that became the central tenet of a new discipline, molecular evolution (2). The most important distinction between the new molecular approach to evolutionary relationships and the older classical ones was that molecules ancestral to a group, whose phenotypes are invariant within the group (i.e., plesiomorphies), could now be used to infer phylogenetic relationships within the group. Thus, by comparing the sequences of molecules whose functions are universal, it was possible not only to construct genealogical trees for Darwin's great kingdoms, but also to go beyond this and construct a *universal* phylogenetic tree, one that united all of the kingdoms into a single phylogenetic "empire."

Ribosomal RNA was central to this endeavor. Not only is the molecule ubiquitous, but it exhibits functional constancy, it changes slowly in sequence, and it is (and was) experimentally very tractable. Moreover, as the central component of the highly complex translation apparatus, rRNA is among the most refractory of molecules to the vagaries of horizontal gene flow, and so was considered likely to avoid the phylogenetic hodgepodge of reticulate evolution and preserve a bona fide organismal trace (3). The rRNA-based universal phylogenetic tree (Fig. 1) brought Biology to an evolutionary milestone, a comprehensive overview of organismal history as well as to the limit of the classical Darwinian perspective.

The initial and strongest impact of the universal tree has been in microbiology. For the first time, microbiology sits within a phylogenetic framework and thereby is becoming a compleat biological discipline: the study of microbial diversity has moved from a collection of isolated vignettes to a meaningful study in relationships. Because niches can now be defined in organismal terms, microbial ecology–long ecology in name only–is becoming ecology in the true sense of the word (7). Yet, the ultimate and perhaps most important impact of the universal phylogenetic tree will be in providing Biology as a whole with a new and powerful perspective, an image that unifies all life through its shared histories and common origin, at the same time emphasizing life's incredible diversity and the overwhelming importance of the microbial world (historically so, and in terms of the biosphere).

A New Era, a New Perspective

In the 1990s, Biology entered the genomic era. It is ironic that (microbial) genomics, which offers such promise for developing the universal phylogenetic tree as a basal evolutionary framework, has seemed initially to do just the opposite. Now that the sequences of many molecules, whose distributions are phylogenetically broad if not universal, are known, biologists find that universal phylogenetic trees inferred from many of them do not fundamentally agree with the rRNA-based universal phylogenetic tree (8). The cause of this incongruity is, of course, reticulate evolution, horizontal gene flow. And the reaction to it–at least according to scientific editorial accounts (9, 10)–has been one of the sky falling. There are grains of truth here. But

*E-mail: carl@ninja.life.uiuc.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

when the scientific sky falls, to mix metaphors, the light dawns. And that is what is now beginning to happen.

The initial reactions to the confusion of trees have been along several lines. One is that the rRNA tree is not the true organismal tree. Unfortunately, no consensus alternative to the rRNA tree emerges from the disparate collection of gene trees that conflict with it; the only concurrence there is, is with the rRNA tree, shown mainly by the componentry of the information processing systems (11), but also more recently by certain whole-genome assessments (12, 13). Another reaction is that the Archaea and Bacteria are specifically related because they have more genes (mainly metabolic) in common with one another than with the eukaryotes (10, 14). This assertion is based on numerology, not phylogenetic analyses; and what it means, frankly, is anybody's guess. In any case, the argument ignores the fact that the phylogenies of the components of the genome replication and expression systems–arguably the most basic systems of the cell–clearly suggest a specific relationship between the Archaea and the eukaryotes, in full agreement with the rRNA tree (11). A third reaction sees horizontal gene transfer as having completely erased any record of the deepest branchings in the universal phylogenetic tree (14–16): the root and earliest branchings of the tree are not knowable. We shall deal with this reaction below.

This confusion and the reaction to it are not because the rRNA tree is somehow wrong (9, 10). An organismal genealogical trace of some kind that goes back in time to the universal ancestor stage does seem to exist (see below), but that trace is carried clearly almost exclusively in the componentry of the cellular information processing systems. The problem here is not with any specific tree or trees, however. We have taken too much for granted about the nature and significance of molecular gene trees; we interpret them from a classical biologist's perspective instead of asking, tabula rasa, what the rRNA (or any other) tree means, what it is telling us about the evolutionary process and about the origin and organization of modern cells.

A Lesson from Some Wanderers. The aminoacyl-tRNA synthetases, perhaps better than any other molecules in the cell, epitomize the current situation and help to understand it. These enzymes have been subject to extensive horizontal gene transfer, from the present day well back into the universal ancestor stage (17–20). Their transfers span the entire phylogenetic gamut, from the species level to transfers between organismal domains (18–20). Their genes tend not to be operonally organized, suggesting that they would generally be transferred independently, which is consistent with the facts (*i*) that the universal phylogenetic trees inferred from these twenty enzymes all differ significantly from one another in various respects, and (*ii*) that only very rarely is some unexpected taxonomic juxtaposition given by more than one of them, an example here being the ostensible sister relationship between the rickettsias and mycobacteria suggested by both the isoleucine and methionine synthetases (20). All twenty or so of these trees differ in various and significant ways from the corresponding rRNA tree, a few radically so (18, 20). But the important point is that, in this set of trees, one can, in the majority of the cases, see a semblance of a common underlying phylogenetic pattern, the same basic branching pattern shown by the rRNA tree (20) (see Fig. 1). [The aminoacyl-tRNA synthetase trees in aggregate also suggest the same major taxonomic groupings within each domain as does the rRNA tree (20)]. This common pattern of the aminoacyl-tRNA synthetase trees cannot itself be the result of horizontal gene transfer (beyond the universal ancestor stage–see below); it is the record of a collective history of these enzymes that has persisted despite horizontal gene transfer. Therefore, the conclusion that an organismal gene trace is preserved in certain of the cell's componentry–a trace that extends back to the stage of the universal ancestor of all extant life–is unavoidable.

Horizontal Gene Transfer

I begin this section with a few simple observations about horizontal gene transfer. For one, transfers can be either selectively driven or selectively neutral. The former have the evolutionary impact, but the latter are simpler to interpret and so should prove the more informative (20). For another, the universality of the genetic code attests to the evolutionary importance of the process. Cells have evolved a number of mechanisms by which to exclude, destroy, or otherwise counteract foreign DNA, much of which is clearly deleterious. Yet none seems to use the strongest defense against alien protein coding genes, i.e., a significantly different genetic code; for if they did, alien genes would be of no value. It seems, therefore, that horizontal gene transfer is not an unavoidable consequence of a universal genetic code, but rather the reverse. Horizontal gene transfer selectively maintains the universality of the genetic code (regardless of how it became established in the first place) because the code is an evolutionary lingua franca required for an essential "genetic commerce" among lineages.

The Evolutionary Roles of Horizontally Acquired and Vertically Generated Variation. Vertically generated and horizontally acquired variation could be viewed as the yin and the yang of the evolutionary process. Without their interplay, evolution as we know it seems impossible. The two are obviously very different in evolutionary impact. Vertically generated variation is necessarily highly restricted in character; it amounts to variations on a lineage's existing cellular themes. Horizontal transfer, on the other hand, can call on the diversity of the entire biosphere, molecules and systems that have evolved under all manner of conditions, in a great variety of different cellular environments. Thus, horizontally derived variation is the major, if not the sole, evolutionary source of true innovation: novel enzymatic pathways, novel membrane transporter capacities, novel energetics, etc.

What then, if anything, is special about vertically generated variation? Is it used simply because it is there and relatively easy to manage? I think not. Vertically generated variation may hold the key to the evolution of biological complexity and specificity. Some time ago I proposed one mechanism whereby this might occur (21), a simple cyclic process that starts with a small homodimeric molecule, the gene for which then undergoes (tandem) duplication, which then allows the homodimer to evolve into a heterodimer (of related subunits), something potentially more sophisticated functionally than the original homodimer. A subsequent mutational event causes the two (tandem) genes to join into a single unbroken reading frame, thereby producing a (symmetric) monomer of approximately twice the original size, the dimerization of which would set the stage for a repeat of the cycle. In this way, a small, simple molecule might evolve into a large, functionally complex one that could have a higher biological specificity and a tighter and more complex coupling to the fabric of the cell. Thus, I would conjecture that the essence of vertically generated variation–variation on a lineage's existing themes–is the principal way in which biological complexity, specificity, and cellular integration evolve. If so, a horizontal acquisition of true novelty and a predominantly vertical generation of complexity, functional differentiation, and integration are the two forces whose interplay propels the evolution of the cell.

Although horizontal transfer and vertical inheritance generally have very different evolutionary consequences, there are conditions–important in the present context–under which their effects mimic one another, indeed become indistinguishable. If organisms A and B are phylogenetically close enough, many of their corresponding proteins differ very little in sequence and not at all in function. Horizontal displacement of a gene for a given protein in organism A by its counterpart from organism B would often have no consequences in terms of the cell, and to the investigator would appear indistinguishable from normal (vertically derived) variation. For this reason alone, neutral horizontal gene transfers of cellular components should predominantly involve closely related species.

How the Organization of the Cell Shapes Horizontal Gene Transfer. One general characteristic of cellular organization in particular strongly affects the quality of horizontal gene flow, namely, the degree and nature of coupling among the various cellular components. Much of cellular componentry is, in effect, modular, i.e., loosely coupled to the cellular fabric. The structure of a modular component is in essence independent of the structure of other components in the cell, and its function is self-defined and minimally connected to other cellular functions. On the other hand, some elements are tightly coupled into the cellular fabric, strongly integrated with others of their kind structurally and/or functionally to make large complexes or complex networks; and they tend for this reason not to be (fully) functional in their own right. The aminoacyl-tRNA synthetases exemplify modular elements, whereas the individual ribosomal proteins exemplify tightly coupled, highly integrated ones. It stands to reason that the genes of modular elements are more readily transferred horizontally than are those of integrated elements. Moreover, transfers of modular elements can easily involve phylogenetically remote donors. In this case the neutral displacement of an element could involve alien and indigenous elements that are significantly different in structure; and examples of this are indeed seen among the aminoacyl-tRNA synthetases (20). However, horizontal displacement involving ribosomal proteins, to the relatively small extent that such occurs, tends to involve only donors closely related to the recipient (unpublished analysis).

Organismal History, Genealogy, and the Concept of Organismal Lineage

What genomics has taught us in no uncertain terms is that the evolutionary history of an organismal lineage is not the same as its genealogy (although in the short term the two may approximate one another). A cell's evolutionary history comprises the histories of all its componentry, from the highly modular to the tightly integrated, at first sight a saga of the reticular wanderings of the modular elements, in which the steady, predictable vertical descent of the tightly integrated elements is lost. Organismal genealogy, on the other hand, is merely the subset of such a history represented in the common evolutionary trace of those components that are central to, integrated into, and so, defining of the fabric of the cell (organism).

What we conventionally take as an organismal genealogical trace is not strictly that. It is the trace of the population to which the individual organism belongs, which makes a subtle but useful change in the way we look at genealogical traces in general and the relationship of horizontal gene transfer thereto. All sexually reproducing species obviously undergo extensive horizontal gene transfer with every generation. But that transfer is strictly confined to the gene pool that comprises the given species. With microorganisms, which basically reproduce asexually, it is much harder to define a comparable gene pool (and so a microbial species). Between closely related microorganisms, extensive genetic transfers, whole chromosomes or large sections thereof, can sporadically occur, as can far more restricted genetic interchanges (down to individual genes), which (latter) are often mediated by various vectors such as phages, plasmids, and naked DNA. However, when the genetic transfers comprise relatively few genes, the donor and recipient organisms need not be phylogenetically close, as dramatically illustrated by a recent transfer of antibiotic resistance (20, 22). Horizontal gene transfer in one form or another occurs at all taxonomic levels. Metaphorically speaking, organismal genealogies are fuzzy lines, fuzzier at some junctures than at others.

Variation acquired horizontally from closely related donors need not significantly blur the organismal genealogical trace, in the limit becoming indistinguishable (in one sense) from vertically generated variation (see above). But "blurring" is also a function of taxonomic level. For example, in the context of phylogenetic relationships among enteric bacteria (*Escherichia*, *Salmonella*, *Proteus*, and so on), the replacement of a ribosomal protein gene in *E. coli* by one from a *Proteus* species would work to blur the organismal trace. However, in the larger taxonomic context of the Proteobacteria as a whole (where *Escherichia*, *Salmonella*, *Proteus*, etc. would generally be lumped together as "the enteric lineage") the same gene transfer would be of no significance. Different taxonomic contexts (levels) require different degrees of "phylogenetic resolution" in the organismal trace. Tracing organismal genealogies is usefully viewed as the tracing of hierarchically nested gene pools, but such pools are obviously ill-defined, context dependent, and not homogeneously mixed. Thus, although a gene may be horizontally transferred, so long as that transfer is basically confined to a natural taxonomic grouping (gene pool), the transferred gene does not erode the genealogical trace of that pool. Organismal lineages may become "fuzzy"–especially as they recede into the deep past–but they are still historically informative of organismal descent so long as their "fuzziness" does not significantly overlap that of other lineages.

What Is the Universal Phylogenetic Tree Telling Us?

The fundamental questions posed by the universal phylogenetic tree are (*i*) the nature of the entity represented by its root and (*ii*) how this entity gave rise to the primary organismal lineages. What stage or stages in the overall history of life on this planet do the ancestor and the emergence of the primary lineages represent? We cannot blithely assume that the universal ancestor is just a typical ancestor, a modern, fully evolved and complex type of cell. At some point in evolutionary history, cells as we know them had to have emerged from some ancient, primitive form of biological organization about which we know nothing. Thus, what the universal ancestor was and how it gave rise to the first lineages are pivotal biological questions.

The time period represented by the primary divergences (branchings) in the universal phylogenetic tree appears to be relatively short compared with that covered by the ramifications within the Bacteria and within the Archaea: Bacteria, at least, have existed for over three billion years, which leaves less than 1.5 billion years for life to pass from the prebiotic chemical stage to the universal ancestor stage and thence to the ancestors of the Bacteria and the other domains (23). Yet, the amount of evolutionary change occurring during the earlier, shorter period far exceeds that occurring during the latter, longer one (23).

Not only was the rate of evolution rapid early on, but it differed in quality from later evolution as well, which can be seen in the sequences of proteins whose distributions are universal: the bacterial version of a universal ribosomal protein tends to be remarkably different from its archaeal equivalent, the same being true, even more dramatically, for the aminoacyl-tRNA synthetases (20, 23). In both cases, in a sequence alignment, a position constant in composition in the Bacteria tends to be so in its archaeal homolog as well, but the archaeal and bacterial compositions for that position often differ from each other. Moreover, among the aminoacyl-tRNA synthetases, a total lack of homology between large (and characteristic) sections of the bacterial version of a molecule and its archaeal counterpart is common (20). Differences of this nature are not the relatively mundane differences one sees among the various Bacteria or among archaeal species. Rather, they appear differences in genre. When eukaryotes are brought into the picture, a characteristic version is again seen, but in almost all cases, the eukaryotic version resembles the archaeal; it exhibits the (general) archaeal genre. In being merely quantitative expressions of divergence, the primary branchings of the universal phylogenetic tree do not convey the fact that they represent major qualitative evolutionary changes, changes whose quality has never been approached in the subsequent evolutionary course.

What evolution in this early period was like can best be sensed from a comparison among the information processing systems. The genome replication mechanisms of Archaea and eukaryotes strongly resemble one another, yet bear no resemblance (in terms of functionally orthologous componentry) to the corresponding bacterial genome replication mechanism. Seemingly, the modern genome replication mechanism has evolved more than once (11). The transcription apparatus of modern cells, on the other hand, does appear to have arisen only once, but during this early period it underwent radical change, major refinement, at least twice—as evidenced by the lack of homology for most of its componentry (except for the main large subunits) between the bacterial version and the archaeal and (closely related) eukaryotic versions (24). Even the translation apparatus underwent significant refinement along these same (phylogenetic) lines during the early period; many ribosomal proteins characteristic of the archaeal and eukaryotic ribosomes have no counterparts on the bacterial ribosome, and vice versa. I take all of this to be a strong indication that the entities represented by the root and earliest branchings of the universal phylogenetic tree were not modern cells, but primitive types of cellular entities in the process of becoming modern cells (23). The universal phylogenetic tree, therefore, is not confined to what we can call the "modern evolutionary era" (once modern cells have come into being). The deepest branchings of this tree take us into uncharted evolutionary waters; the door to understanding earlier, more primitive forms of life has opened.

Cellular Evolution

Modern cells are fully evolved entities. They are sufficiently complex, integrated, and "individualized" that further major change in their designs does not appear possible, which is not to say that relatively minor (but still functionally significant) variations on existing cellular themes cannot occur or that, under certain conditions, cellular design cannot degenerate. On Earth today, there exist three distinct cellular designs: the bacterial, the archaeal, and the eukaryotic (4, 23, 25). In their (localized) evolutionary wanderings, the gene pools of these three types no longer come into major contact with one another; occasional horizontal exchanges of certain modular elements do, of course, occur. The universal phylogenetic tree takes us back to an era before any of this had happened.

If ever there had been a stage in the history of life when cells were simpler in design than they now are–i.e., had less complex, less integrated, more modular componentry–horizontal gene transfer would then have been a more dominant evolutionary force than it now is. And in the transition from an abiotic, chemically reactive planet Earth to the living forms with which we are familiar, there reasonably had to be such a stage or stages. As I have argued previously from a somewhat different perspective (26), when cells are simple enough, horizontal gene transfer totally dominates the evolutionary scene, and all life becomes a single, diverse gene pool; all of the cell's componentry can be subject to horizontal gene flow. It is only in this way that the radical novelty needed to progressively boot-strap primitive cellular entities into modern cells can occur. I do not imply here that vertical inheritance does (or did) not exist at such stages. Vertical inheritance, implicit in the process of cellular replication, had to exist, but at the universal ancestral stage it lead to cell lines, gene pools, that rapidly (on the evolutionary time scale) turned over: any specific organismal trace that they possessed was sooner or later washed away by horizontal gene flow. At such a stage, evolution was in effect communal: there was a progressive evolution of the whole, not an evolution of individual organismal lineages *per se* (26–28).

In this evolutionary process a stage inevitably will be reached where some cellular entities become complex enough that their cell designs start to become unique: different ones of them take their emerging cellular designs down different evolutionary avenues. In other words, from the universal gene pool (the communal ancestor), there will emerge "individual" pools, characterized by the fact that horizontal gene transfer continues to occur in a more or less rampant fashion within each, but between which horizontal gene flow becomes progressively restricted in character, because the universal systems of the cell are becoming increasingly complex and idiosyncratic in each separate pool, and new components are evolving in each of them that have little or no functional significance in any other pool.

If this picture of cellular evolution is correct, there will come a point at which certain of the cell's componentry becomes sufficiently complex (idiosyncratic) and sufficiently integrated into the emerging cellular fabric that horizontal gene displacement (especially of the phylogenetically distant variety) will not strongly influence them. This point should be reached first for the complex components most central to, and most defining of, the cellular fabric (26). (Gradually, others of the cell's componentry will then follow suit, becoming more or less refractory to horizontal gene flow.) Then cells will have reached a stage where ephemeral cell lines give way to stable cellular lineages, where true organismal genealogies can arise. *The initial bifurcation in the universal phylogenetic tree marks this point* (Fig. 1).

At first the organismal trace is carried by only a few molecules, those most important to, and most defining of, the cell's design (26). These primary organismal lineages will be of the "fuzzy" variety (see above), for whereas the cellular subsystems carrying the trace are more or less refractory to horizontal displacement in general, they are not altogether refractory to displacement from within their own gene pool. Only when immunity to displacement extends to the specific pools themselves–only when the individual pools in turn spawn subpools and so on–will these subsystems become fully locked-in, subject to change basically through vertical inheritance. The balance between horizontally acquired and vertically generated variation will continue to change until the evolution of the cell is complete, until the complex (finalized) modern cell types emerge.

The above theory makes a testable prediction: the ancestors of the individual domains—the Bacteria, the Archaea, and the eukaryotes—are each communal, and the evidence for their communal nature, in the form of elevated levels of horizontal gene transfer within each domain early on (i.e., transfer involving the ancestors of the major taxa), should still exist. Sufficient sequence data do not now exist to test this prediction definitively. Existing data, however, are consistent with it (20).

Summary and Conclusion

The universal phylogenetic tree based on rRNA is a valid representation of organismal genealogy. But it is unlike any other phylogenetic tree. It transcends the era of modern cells; its deepest branchings extend back in time to an era when cellular entities were considerably more primitive than cells are today. These primitive entities were basically modular (loosely coupled) in construction. They were not highly integrated modern cells (which, however, still contain a good deal of modular componentry). And they consequently engaged in a rampant form of horizontal gene transfer, transfers that jumbled their histories. The primary bifurcation of the universal phylogenetic tree represents the first evolutionary stage at which cellular design became sufficiently stable that horizontal gene transfer could not completely wash away a collective (organismal) trace, and true organismal lineages then gradually began to consolidate.

At first the organismal trace resided in a few molecular species only, those that had become sufficiently complex (tightly coupled) and woven into the emerging cellular fabric that they were largely refractory to global horizontal gene displacement. Chief among these components was the primitive translation apparatus, especially its RNA component. Gradually, as cells became increasingly integrated entities, other cellular functions followed suit (became more or less refractory to horizontal gene flow). However, still others of them remained, and remain today, subject to the vagaries of horizontal gene flow.

Much of this picture is captured in the evolutions of the aminoacyl-tRNA synthetases, modular components of the translation apparatus that are subject to widespread horizontal gene transfer (18–20). Their aboriginal evolutionary histories have been severely jumbled by horizontal gene flow, yet, in the aggregate, their phylogenetic trees retain clear vestiges of the ground structure of the universal tree. Among the cell's components that are modular to one extent or another are metabolic enzymes and pathways, transmembrane proteins, and much of cellular energy metabolism. And through genomics we have begun to understand the horizontal wanderings of at least some of these (29, 30).

The high, pervasive levels of horizontal gene transfer at early times created an evolutionarily communal state of living systems in the sense that the aboriginal organismal community evolved as a collective whole, not as individual cellular lineages. With the inevitable emergence of complexity in early cellular entities, a stage was reached where distinct subpopulations emerged from the universal ancestral communal state. Each was distinguished by the fact that horizontal gene transfer continued to be pervasive within it (the communal state persisted), but between separate pools the level and scope of genetic exchange diminished, because in each of them different, incompatible cell designs were being worked out. Horizontal gene transfer continued to abate within and between these pools–three of which would refine into the ancestors of the three extant organismal domains–until the evolution of the cell (in each) attained a (modern) fully evolved state. This stage was probably reached as the major lineages in each organismal domain emerged.

The universal phylogenetic tree opens the doors to past and future, to the two greatest challenges the biologist faces: on the one hand, the evolution of the cell, the challenge of reconstructing our biological past; on the other, the nature of the biosphere, comprehending our future, understanding the evolutionary and other interactions (mainly) among microorganisms that form and sustain this living planet. The 19th century laid the three great foundation stones of Biology: cell theory, Mendelian genetics, and Darwinian evolution. The spectacular biological edifice of the 20th century, molecular biology, was built on the first two of these. But the universal phylogenetic tree shows us that the edifice that is 21st Century Biology will rest solidly on all three.

The author is most appreciative to Norman Pace and Charles Kurland for their readings of this manuscript and the very helpful suggestions that have come therefrom. This work is supported by a grant from the National Aeronautics and Space Administration.

1. Darwin, C. (1857) (1: p. 456) in a letter to T. H. Huxley (Corresp. 6: p.456) (1887) in *The Life and Letters of Charles Darwin, Including an Autobiographical Chapter*, ed. Darwin, F. (John Murray, London), Vol. 2, p. 456.
2. Zuckerkandl, E. & Pauling, L. (1965) *J. Theor. Biol.* **8,** 357–366.
3. Fox, G. E., Pechman, K. R. & Woese, C. R. (1977*) Int. J. Syst. Bacteriol.* **27,** 44–57.
4. Woese, C. R., Kandler, O. & Wheelis, M. L. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 4576–4579.
5. Fox, G. E., Stackebrandt, E., Hespell, R. B., Gibson, J., Maniloff, J., Dyer, T. A., Wolfe, R. S., Balch, W. E., Tanner, R. S., Magrum, L. J., *et. al*. (1980) *Science* **209,** 457–463.
6. Iwabe, N., Kuma, K., Hasegawa, M., Osawa, S. & Miyata, T. (1989) *Proc. Natl. Acad. Sci.* USA **86,** 9355–9359.
7. Olsen, G. J., Lane, D. J., Giovannoni, S. J., Pace, N. R. & Stahl, D. A. (1986) *Annu. Rev. Microbiol.* **40,** 337–365.
8. Brown, J. R. & Doolittle, W. F. (1997) *Microbiol. Mol. Biol. Rev.* **61,** 456–502.
9. Pennisi, E. (1998) *Science* **280,** 672–674.
10. Pennisi, E. (1999) *Science* **284,** 1305–1307.
11. Olsen, G. J. & Woese, C. R. (1996) *Trends Genet.* **12,** 377–379.
12. Snel, B., Bork, P. & Huynen, M. A. (1999) *Nat. Genet.* **21,** 108–110.
13. Fitz-Gibbon, S. T. & House, C. H. (1999) *Nucleic Acids Res.* **27,** 4218–4222.
14. Philippe, H. & Forterre, P. (1999) *BioEssays* **21,** 871–879.
15. Doolittle, W. F. (1999) *Science* **284,** 2124–2128.
16. Martin, W. (1999) *BioEssays* **21,** 99–104.
17. Nagel, G. M. & Doolittle, R. F. (1995) *J. Mol. Evol.* **40,** 487–498.
18. Brown, J. R. (1998) in *Thermophiles: The Keys to Molecular Evolution and the Origin of Life?* eds. Wiegel, J. & Adams, M. H. W. (Taylor & Francis, London).
19. Wolf, Y. I., Aravind, L., Grishin, N. V. & Koonin, E. V. (1999) *Genome Res.* **9,** 689–710.
20. Woese, C. R., Olsen, G. J., Ibba, M. & Soll, D. (2000) *Microbiol. Mol. Biol. Rev.* **64,** 202–236.
21. Woese, C. R. (1971) *J. Theor. Biol.* **33,** 29–34.
22. Brown, J. R., Zhang, J. & Hodgson, J. E. (1998) *Curr. Biol.* **8,** R365–R367.
23. Woese, C. R. (1987) *Microbiol. Rev.* **51,** 221–271.
24. Langer, D., Hain, J., Thuriaux, P. & Zillig, W. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 5768–5772.
25. Graham, D., Overbeek, R., Olsen, G. J. & Woese, C. R. (2000) *Proc. Natl. Acad. Sci.* USA **92,** 3304–3308.
26. Woese, C. (1998*) Proc. Natl. Acad. Sci.* USA **95**, 6854–6859.
27. Woese, C. R. (1982) *Zentralbl. Bakteriol. Mikrobiol. Hyg. Abt. 1 Orig. C* **3,** 1–17.
28. Kandler, O. (1994) *J. Biol. Phys.* **20,** 165–169.
29. Bult, C. J., White, O., Olsen, G. J., Zhou, L., Fleischmann, R. D., Sutton, G. G., Blake, J. A., FitzGerald, L. M., Clayton, R. A., Gocayne, J. D., *et al*. (1996) *Science* **273,** 1058–1073.
30. Kawarabayasi, Y., Sawada, M., Horikawa, H., Haikawa, Y., Hino, Y., Yamamoto, S., Sekine, M., Baba, S., Kosugi, H., Hosoyama, A., *et al*. (1998) *DNA Res.* **5,** 55–76.